Short Communication

Phonological processing in speech perception: What do sonority differences tell us?

CrossMark

Isabelle Deschamps [b,d,e,*], Shari R. Baum [a,b], Vincent L. Gracco [a,b,c]

[a] McGill University, Faculty of Medicine, School of Communication Sciences and Disorders, 1266 Avenue des Pins, Montreal, Quebec H3G 1A8, Canada
[b] Centre for Research on Brain, Language and Music, Rabinovitch House, McGill University, 3640 rue de la Montagne, Montreal, Quebec H3G 2A8, Canada
[c] Haskins Laboratories, 300 George St., Suite 900, New Haven, CT 06511, USA
[d] Rehabilitation Department, Laval University, Quebec, QC, Canada
[e] Centre de Recherche de l'Institut Universitaire en santé mentale de Québec, Quebec, QC, Canada

ARTICLE INFO

ABSTRACT

Previous research has associated the inferior frontal and posterior temporal brain regions with a number of phonological processes. In order to identify how these specific brain regions contribute to phonological processing, we manipulated subsyllabic phonological complexity and stimulus modality during speech perception using fMRI. Subjects passively attended to visual or auditory pseudowords. Similar to previous studies, a bilateral network of cortical regions was recruited during the presentation of visual and auditory stimuli. Moreover, pseudowords recruited a similar network of regions as words and letters. Few regions in the whole-brain results revealed neural processing differences associated with phonological complexity independent of modality of presentation. In an ROI analysis, the only region sensitive to phonological complexity was the posterior part of the inferior frontal gyrus (IFGpo), with the complexity effect only present for print. In sum, the sensitivity of phonological brain areas depends on the modality of stimulus presentation and task demands.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Neuroimaging studies have identified a distributed network of brain regions involved in the processing of phonological information during the perception of speech sounds. For example, studies that have investigated phonological processing by contrasting the processing of syllables or phonemes to the processing of complex auditory stimuli (e.g. environmental sounds (Giraud & Price, 2001), bird songs (Tremblay, Baroni, & Hasson, 2012), tones (Demonet et al., 1992; Poeppel et al., 2004; Rimol, Specht, & Hugdahl, 2006; Vouloumanos, Kiehl, Werker, & Liddle, 2001), and unintelligible speech sounds (Benson, Richardson, Whalen, & Lai, 2006; Liebenthal, Binder, Spitzer, Possing, & Medler, 2005; Obleser, Zimmermann, Van Meter, & Rauschecker, 2007; Okada et al., 2010)) have consistently reported clusters of activation within the superior temporal gyrus (STG) and superior temporal sulcus (STS). In addition, the presentation of auditory and/or orthographic stimuli (word and/or pseudowords) requiring a phonological judgment recruits regions located within the inferior frontal gyrus (IFG), the middle frontal gyrus (MFG) and the inferior parietal lobules (IPL) (Booth et al., 2002; Burton, Locasto, Krebs-Noble, & Gullapalli, 2005; Burton, Small, & Blumstein, 2000; Jacquemot, Pallier, LeBihan, Dehaene, & Dupoux, 2003; Kareken, Lowe, Chen, Lurito, & Mathews, 2000; Poldrack et al., 2001). Of particular interest is the observation that the IFG, MFG and IPL are typically not recruited during passive listening (Deschamps & Tremblay, 2014; McGettigan et al., 2011; Tremblay & Small, 2011) or passive reading of single letters (Van Atteveldt, Formisano, Goebel, & Blomert, 2004). For example, studies that have used a passive paradigm to examine specific phonological processes such as print-speech convergence of letters (Van Atteveldt et al., 2004), or manipulated phonological complexity by comparing single consonants to consonant clusters in words (Tremblay & Small, 2011), pseudowords (McGettigan et al., 2011) and syllable sequences (Deschamps & Tremblay, 2014) during passive listening have not reported clusters of activation within the IFG, the MFG and the IPL (Deschamps & Tremblay, 2014; McGettigan et al., 2011; Tremblay & Small, 2011; Van Atteveldt et al., 2004). In fact, Deschamps and Tremblay (2014) and Tremblay and Small (2011) reported that activation magnitude increased as a function of syllabic complexity with auditory stimuli

* Corresponding author at: Centre de Recherche de l'Institut Universitaire en santé mentale de Québec (CRIUMSQ) 2601, chemin de la Canardière, office: F-2424A, Québec, Québec G1J 2G3, Canada.
   E-mail address: isabelle.deschamps.1@ulaval.ca (I. Deschamps).

only within superior temporal plane regions, whereas McGettigan et al. (2011) did not find any positive correlations between syllabic complexity and brain activation. In addition, Van Atteveldt et al. (2004) found that the passive viewing of letters activated the bilateral lateral and inferior occipital cortex, and the passive listening to single speech sounds activated the bilateral primary auditory cortex, the STG and the STS. Interestingly, the bilateral posterior STS and STG were activated during both the passive viewing of letters and the passive listening to speech sounds, suggesting that this region might be involved in heteromodal phonological processing (i.e. multisensory convergence).

Logically, measuring brain activation when subjects are passively attending to stimuli can target more automatic and obligatory phonological processes, whereas more cognitively demanding tasks (e.g. discrimination, rhyme judgments) recruit additional processes (i.e. verbal working memory, segmentation, rehearsal). While passive processing of spoken or printed letters and words in the absence of a task does activate a number of auditory and visual areas, as well as regions involved in phonological processing, it is not clear whether the same is true of pseudowords. One study by Burton et al. (2005) using two different tasks (i.e. rhyming and same/different judgments) with auditory and visually presented words and pseudowords identified regions that were modality-specific (i.e. left STG for auditory stimuli and right lingual gyrus for visual stimuli) and a number of regions that were recruited across modalities, stimulus type (i.e. words and pseudowords) and tasks (e.g. left IFG, bilateral posterior STG, left fusiform). Of particular interest is that Burton et al. (2005) found more activation across brain regions for pseudoword judgments and interpreted the result as reflecting greater articulatory recoding demands. In fact, differences between the processing of words and pseudowords during tasks requiring a judgment and reading are well documented (Mechelli, Gorno-Tempini, & Price, 2003; Shaul, Arzouan, & Goldstein, 2012; Simos et al., 2002; Xiao et al., 2005). Notwithstanding the importance of these results, the findings raise the issue of whether the brain regions outside of STS/STG were recruited because of the task or because of the pseudowords. This is a potentially important issue in that pseudowords are often used to evaluate reading ability differences in fluent and dysfluent readers (Shaul et al., 2012) and in children with specific speech and language impairments (Macchi, Schelstraete, & Casalis, 2014). In order to investigate more automatic phonological processes, we used a passive print and speech pseudoword paradigm. By minimizing the influence of non-linguistic cognitive functions and semantic/conceptual activation, the neural correlates associated with automatic modality-specific or heteromodal phonological processes were investigated.

In the current study we used a metric of phonological complexity to examine obligatory phonological processing targeting the structure of the syllable. Because phonological complexity reflects more than simple speech/non-speech distinctions, differences in stimulus length, or the presence or absence of consonant clusters, we focused on sonority. Sonority has been used to explain a wide range of linguistic and psycholinguistic phenomena (Bastiaanse, Gilbers, & Van Der Linde, 1994; Clements, Kingston, & Beckman, 1990; Goad, 2010; Romani & Galluzzi, 2005; Zec, 1995). In this study, we varied the sonority values of consonants within the initial consonant cluster (i.e. CC) of the first syllable in pseudowords. The principle of sonority stipulates that speech sounds can be characterized according to their placement along a scale (Clements et al., 1990; Steriade, 1990) that captures the relative resonance of speech sounds (Clements, 2009). For example, in English, vowels are the most sonorous and stop consonants are the least sonorous (Dobrovolsky & Katamba, 1996). Since each consonant and vowel has a sonority value, sonority differences between two phonemes (in our case consonants) can be calculated (Gierut, 2007). In a

consonant cluster, the smaller the difference between the sonority of two consonants, the more phonologically complex the cluster (for more details, refer to Gierut, 2007; Steriade, 1990). While, the influence of sonority on speech perception, speech production and reading has been documented behaviorally during language acquisition (first or second) and in neuropsychological populations (Baum, 2002; Fabre & Bedoin, 2003; Miozzo & Buchwald, 2013; Morrisette, Farris, & Gierut, 2006; Romani & Calabrese, 1998; Sperbeck & Strange, 2010), the neural correlates underlying the processing of sonority, to our knowledge, have not yet been investigated.

Informed by the results from previous studies, we expected to find a subset of regions located within the superior temporal plane and STS that have been identified by prior neuroimaging studies as relevant to phonological processing that would be sensitive to syllabic complexity while subjects are passively attending to visually- and/or auditorily-presented pseudowords; such a finding would highlight regions that are involved in obligatory and mandatory phonological processes and distinguish modality-specific from heteromodal phonological processes.

2. Material and methods

2.1. Participants

Eighteen healthy right-handed (Oldfield, 1971) native speakers of Canadian English participated in this experiment (mean age 24 years + 7.7, range: 18–40 years, 11 females). The data from two participants could not be used due to technical problems during the acquisition of the high-resolution anatomical scan, leaving sixteen participants in the analysis. All participants had normal hearing and normal or corrected-to-normal vision. Subjects had no self-reported history of speech, language or neurological disorders. All participants gave informed consent in accordance with the ethics committee of the Montreal Neurological Institute (MNI). The study was approved by the Magnetic Resonance Research Committee (MRRC) and the MNI Research Ethics Board.

2.2. Stimuli

The experiment consisted of four tasks: (1) passive listening, (2) passive reading, (3) listening and repeating, and (4) reading aloud. Only the first two tasks (i.e. passive listening and passive reading) were analyzed for the current report. The stimuli consisted of a set of 40 pseudowords presented visually and auditorily. To create pseudowords, initial word lists containing common two-syllable trochaic nouns, six to eight letters in length, with onsets utilizing all legal two-consonant clusters of English were developed using the UWA Psychology: MRC Psycholinguistic Database (Colheart, 1981). Words were ordered by sonority ranking (Steriade, 1990) and divided into two categories on that basis: low phonological complexity (LPC) and high phonological complexity (HPC). We calculated the sonority ranking by measuring the absolute distance in sonority between the two consonants in each word-onset cluster. Words with consonant cluster onsets that had a sonority ranking of 4 or more (e.g. /pl/) were classified as LPC and words with consonant cluster onsets that had a sonority ranking of 3 or less were classified as HPC (e.g. /st/). Based on these words, pseudowords were then created by substituting the first consonant of the onset of the second syllable for another English consonant (see Supplementary material S1 for some examples). Pseudowords were compared for bigram frequency of the first and second consonant of the onset of the second syllable to ensure that they were legal and equally frequent combinations of English orthography and phonology (Balota et al., 2007; Solso & Juel, 1980).

Pseudowords in each phonological category were matched in terms of number of orthographic neighbors and bigram frequency mean (Balota et al., 2007). The auditory pseudowords were recorded by a female native English speaker in a sound-treated room. Recordings were made at a 44 kHz sampling rate directly onto disk.

2.3. Procedure

The experiment consisted of 4 experimental runs (13.5 min each) that included two perceptual runs and two production runs. Within each run, forty stimuli were presented in the auditory modality and another forty in the visual modality, for a total of 80 stimuli per run. The stimulus modality was randomized within each run. The perceptual runs were always presented first in order to minimize the likelihood of covert rehearsal. Participants were not aware until the beginning of the production runs when they would have to speak in the scanner. The perceptual runs were separated from the production runs by the acquisition of a high-resolution anatomical scan. For both the perceptual and production runs, auditory stimuli were on average 1000 ms in duration and the visual stimuli remained on the screen for 1000 ms. Both stimulus presentation and participants' responses occurred during the interscan interval (i.e. silent interval). All stimuli were presented using Presentation Software (Neurobehavioral System, CA, USA). A resting condition signaled by a dark gray screen was also included as the baseline condition. In order to verify that subjects were paying attention during the perceptual and production runs, catch trials were included. Upon visual presentation of a fixation cross, subjects were instructed to press a button on an MRI-compatible response box for the catch trials. Rest trials were interleaved with the experimental conditions. Within each run, the number of rest trials (10) was optimized using OPTseq2 (http://surfer.nmr.mgh.harvard.edu/optseq/).

2.4. fMRI parameters

The data were acquired on a 3T Siemens Trio scanner at the Montreal Neurological Institute. Participants wore MR-compatible headphones (Sensimetrics Corporation, Malden, MA) and their head was immobilized by means of a vacuum-bag filled with polystyrene balls and a forehead-restraining device (Hybex Innovations, St-Leonard, Qc, CAN). A T2-weighted gradient-echo multi-slice EPI interleaved sequence was used for the fMRI scans (TE = 30 ms, TR = 2.04, Flip Angle 90°, matrix 64 × 64, FOV = 256 × 256, slice thickness 4 mm, isotropic, no gap). Thirty-four axial slices oriented parallel to the AC–PC line were acquired covering the whole brain. To eliminate movement artefact associated with speaking and to ensure that subjects could clearly hear the auditory stimuli, a clustered sparse temporal acquisition paradigm was used. For each trial, a clustered acquisition of two volumes was completed, resulting in 208 functional volumes per experimental run. The silent inter-scan interval was 4.04s (cluster-onset asynchrony: 8.08s). Stimulus presentation started during the inter-scan interval exactly 4 s before the acquisition of the two successive volume scans. High-resolution T1-weighted volumes were acquired for anatomical localization after the two perceptual runs.

2.5. fMRI data analysis

The four time series were spatially registered, motion-corrected (within and across runs), de-spiked and converted to a percentage of signal change using AFNI (Cox, 1996). The anatomical scan of each participant was aligned to their registered EPI time series using local Pearson correlations (Saad et al., 2009). A linear least squares model was used to fit to each time point of the hemodynamic response function for each of the conditions. Each experimental condition had its own regressor. Additional regressors for the mean, the linear and the quadratic trend components as well as the six motion parameters (x, y, z, roll, pitch, yaw) were also included in the model. We modeled a 2.02s period beginning at the start of the stimulus, using AFNI's TENT function. We used the first TR for all subsequent analyses. To create a surface representation of each participant's anatomy, Freesurfer was used (Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999). For each participant, each hemisphere of the anatomical volume was inflated to a surface representation and aligned to a template of average curvature. SUMA was used to import the surface representation from Freesurfer and to project the functional data from the first level analysis onto the 2D surface. Both the surface representations and the functional data were standardized to a common mesh reference system (Saad, Reynolds, Argall, Japee, & Cox, 2004). The functional data were smoothed on the surface using a Gaussian 6-mm full-width at half-maximum filter. Whole-brain group analyses were performed on the surface using SUMA on the subjects' beta values taken from the first level analysis. The main focus of the whole-brain analyses was on the effect of auditory and visual stimuli during speech perception as well as the effect of phonological complexity during speech perception. The surface-based group analyses were corrected for multiple comparisons, using a Monte Carlo simulation implemented in Freesurfer. This correction implements a cluster-size threshold procedure to protect against Type I error. Based on the simulation results, it was determined that a family-wise error (FEW) rate of $p < 0.01$ is achieved with a minimum cluster size of 127 contiguous surface nodes, each significant at $p < 0.01$. From the whole-brain contrasts (corrected for multiple comparisons) we also identified brain areas that were sensitive to both auditory and visual stimuli (auditory ∩ visual) using a conjunction mask of brain activity (Nichols, Brett, Andersson, Wager, & Poline, 2005).

2.6. Anatomical ROI analysis

To further profile the role of regions typically reported in studies on phonological processing (Burton et al., 2005; Okada & Hickok, 2006; Price, 2012; Vouloumanos et al., 2001), we conducted an analysis of anatomical regions of interest (ROI). This subset of ROIs included the inferior frontal gyrus (pars opercularis (IFGpo), pars triangularis (IFGpt), and pars orbitalis (IFGporb), the planum temporale (PT), the superior temporal sulcus (STS), the lateral superior temporal gyrus (STG), and the supramarginal gyrus (SMG). Each ROI was defined on the subject's individual cortical surface representation using an automated parcellation scheme (Desikan et al., 2006; Fischl et al., 2004). This parcellation scheme relies on a probabilistic labeling algorithm based on the anatomical convention of Duvernoy (1991) (Destrieux, Fischl, Dale, & Halgren, 2010). For some of the ROIs selected, we edited the Freesurfer parcellation by sub-dividing it into smaller ROIs (See supplementary S2). Details of the parcellation are described in Supplementary Materials S2.

For each subject, we extracted the mean percentage of BOLD signal change in each ROI. We first examined which ROIs were significantly active in perception by testing the following hypothesis using FDR-corrected t-tests (Benjamini & Hochberg, 1995; Genovese, Lazar, & Nichols, 2002) (q = 0.05): (i) perception > 0, (n = 16, one-sample t-tests). All the ROIs that were significantly active were submitted to statistical evaluation in a repeated-measures ANOVA (rANOVA) with the factors modality (auditory and visual) and complexity (high phonological complexity and low phonological complexity). We investigated the main

effects of modality and complexity as well as two-way interactions.

3. Results

3.1. fMRI results

3.1.1. Whole brain analyses

The node-wise ANOVA showed a significant effect for the auditory and visual modality during speech perception (Fig. 1A). The activation associated with the auditory and orthographic stimuli revealed regions involved in the sensory processing of auditory or visual information, namely bilateral primary visual cortex and its corresponding association areas (e.g. lingual gyrus, fusiform gyrus, middle occipital gyrus and inferior occipital sulcus and gyrus) when pseudowords were presented orthographically and bilateral primary auditory cortex and its association areas (e.g. lateral superior temporal gyrus, transverse temporal gyrus and sulcus) when pseudowords were presented auditorily (for a review, Price, 2012). As illustrated in Fig. 1A, the conjunction between the auditory and visual stimuli revealed overlapping activation for both modalities in numerous bilateral cortical regions including the posterior portion of the STS, the inferior circular sulcus of the insula, the posterior cingulate gyrus and sulcus, the calcarine sulcus, and the medial superior frontal gyrus. Overlapping clusters of activation were also found in the left SMG, the left superior frontal gyrus and sulcus, the left middle frontal gyrus, the left postcentral gyrus, the left precentral gyrus, the left mid-portion of the STG, the left posterior portion of the inferior temporal gyrus, the left cuneus, the left precuneus, the right superior parietal gyrus, and the right anterior portion of the cingulate gyrus and sulcus. The

node-wise ANOVA showed a significant effect of complexity for speech perception (Fig. 1B). When we compared the HPC and LPC activation to investigate regions sensitive to phonological complexity, significant differences were observed in the left calcarine sulcus, middle occipital gyrus, occipital sulcus and STS as well as the right occipital pole and calcarine sulcus (see Fig. 1B, and Table 1 for a complete list).

3.1.2. ROI analysis

Only the ROIs that were significantly activated for speech perception were included in the subsequent analyses. Eight ROIs (bilateral IFGpo, PT, STGp, left SMGa and right SMGp) were used to investigate the main effect of modality (i.e. auditory, visual), the main effect of phonology (HPC, LPC), and the two-way interaction between modality $*$ phonology.

3.1.2.1. Main effects. A main effect of modality was found in the bilateral PT (left PT: $F_{1,15} = 36.22$, $p < 0.0001$; right PT: $F_{1,15} = 44.64$, $p < 0.0001$) and bilateral STGp (left STGp: $F_{1,15} = 34.60$, $p < 0.0001$; right STGp: $F_{1,15} = 85.10$, $p < 0.0001$). Paired sample t-tests revealed that both of these regions were significantly more active for auditory stimuli relative to visual stimuli (left PT: $t = 6.02$, $p < 0.0001$; right PT: $t = 6.68$, $p < 0.0001$; left STGp: $t = 5.88$, $p < 0.0001$; right STGp: $t = 9.23$, $p < 0.0001$). To determine whether the difference observed was due to a lack of activation in one modality, we tested whether the activation level in each modality was significantly different from 0 (one-sample t-test, one-tailed). The left PT was significantly activated for both modalities (auditory: $t = 7.44$, $p < 0.0001$, visual: $t = 2.1$, $p = 0.026$), while the bilateral STGp (left STGp auditory: $t = 6.2$, $p < 0.001$, left STGp visual: $t = 1.19$, $p = 0.28$; right STGp auditory:
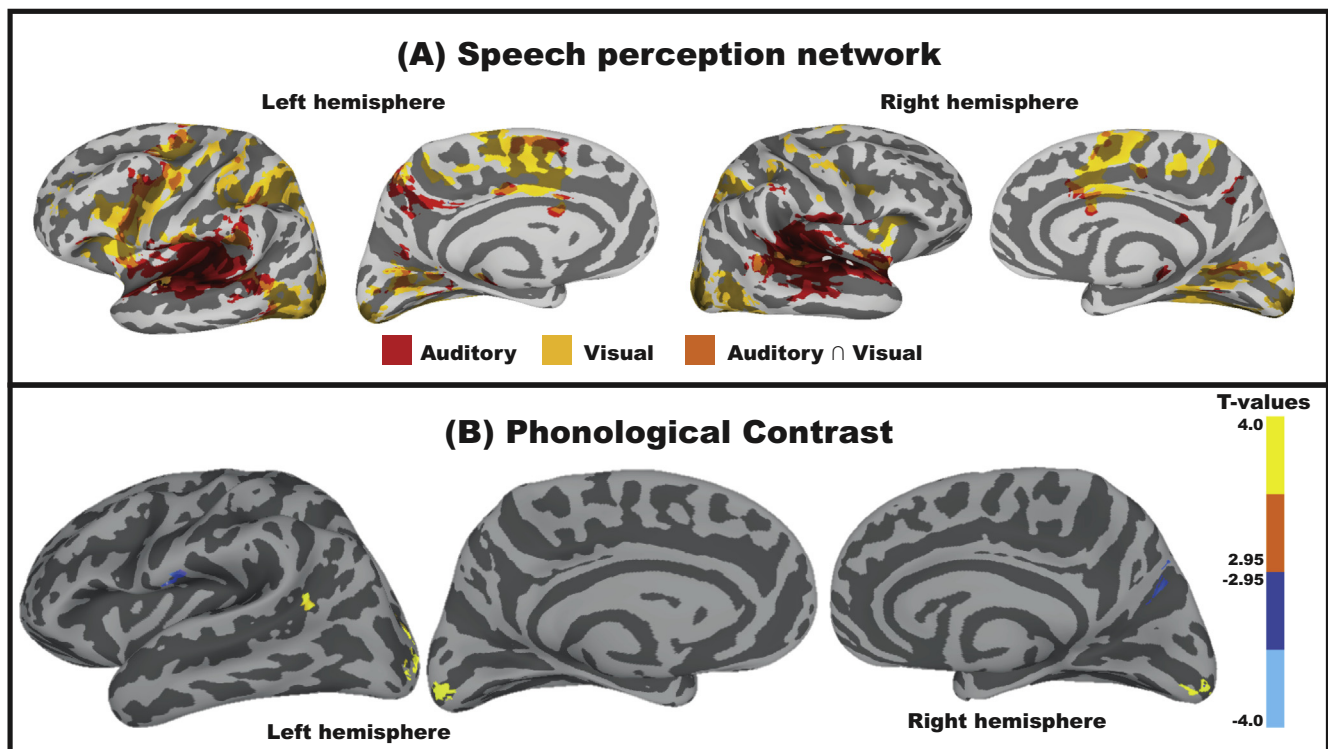

Fig. 1. Part A. Speech perception network. The red color scheme represents nodes that are significantly active during the auditory condition (i.e. passive listening), the yellow color scheme represents nodes that are significantly active during the visual condition (i.e. passive reading) and the orange color scheme represents nodes that are significantly active for the conjunction of the two conditions (i.e. passive listening and passive reading). Part B. Phonological contrast. Clusters of significant differences between the HPC and LPC conditions. Positive activation is represented in yellow and negative activation is represented in blue. Activation is shown on the group average smoothed white matter folded surface.

Table 1
FWE-corrected group-level (N = 16), whole brain results for the contrasts between HPC and LPC. Coordinates are in Talairach and represent the peak surface for each node of the cluster (minimum cluster size: 127 contiguous surface nodes, each significant at $p < 0.01$).

Description	Hemi	X	Y	Z	t	p	Nodes	Area
Calcarine sulcus	Left	−7	−94	3	4.02	0.001	252	116.36
Middle occipital gyrus, occipital pole, middle occipital sulcus, and superior occipital sulcus		−26	−98	−7	3.46	0.003	298	108.56
Superior temporal sulcus		−42	−51	16	3.47	0.003	145	49.47
Subcentral gyrus and sulcus		−62	−11	14	−3.04	0.008	155	46.57
Occipital pole and calcarine sulcus	Right	17	−100	−6	4.44	0.0004	140	65.56
Parietal–occipital sulcus		15	−67	35	−3.35	0.004	128	33.76

$t = 9.4$, $p < 0.0001$; right STGp visual: $t = 0.34$, $p = 0.74$) and the right PT (auditory: $t = 6.22$, $p < .0001$; visual $t = 0.87$, $p = 0.40$) were not significantly activated in the visual condition (for more details, refer to Fig. 2A).

3.1.2.2. Two-way interaction effects (Phonology × Modality). A two-way interaction was found for the bilateral IFGpo (left IFGpo: $F_{1,15} = 5.11$, $p = 0.04$; right IFGpo: $F_{1,15} = 5.62$, $p = 0.03$), the left SMGa ($F_{1,15} = 5.77$, $p = 0.03$) and the right SMGp ($F_{1,15} = 8.461$, $p = 0.01$). Paired sample t-tests revealed a significant HPC–LPC difference in the visual modality ($t_{15} = 2.368$, $p = 0.03$) for the left IFGpo (refer to Fig. 2B). A significant difference between visual and auditory modalities for the HPC–LPC contrast emerged in the right IFGpo ($t_{15} = 2.371$, $p = 0.03$) and right SMGp ($t_{15} = 2.909$, $p = 0.01$). For the left SMGa, paired sample t-tests revealed a marginally significant effect of complexity only in the visual modality ($t_{15} = 2.127$, $p = 0.05$).

4. Discussion

The findings from the current experiment demonstrate that in the absence of an explicit task, only a subset of regions typically involved in phonological processing are sensitive to sonority differences that modulate syllabic complexity. We focused on a subset of brain regions within the posterior superior temporal plane and the inferior frontal cortex and examined brain regions associated with the manipulation of phonological complexity (sonority) and stimulus modality (orthographic and auditory) in pseudowords. The results suggest that neural processing differences associated with phonological complexity during passive listening are modality dependent. In the following, we discuss the findings of the whole-brain analyses and ROI analyses in terms of the role of phonological complexity and modality of presentation as an experimental tool to flesh out the neural correlates of phonological processing.

Not surprisingly, the whole-brain phonological contrast yielded few regions in which an effect of complexity was observed,

suggesting that in the absence of an overt task, regions typically involved in the processing of phonological information are not modulated by complexity. This finding is congruent with two recent neuroimaging studies in which phonological complexity was manipulated during passive listening. In these studies, phonological complexity was manipulated by contrasting consonant clusters (CCV) to single consonant vowel combination (CV) in words (Tremblay & Small, 2011) and pseudowords (McGettigan et al., 2011). In the first study, the phonological contrast during passive perception revealed one region within the right PT in which the activity was scaled to the degree of complexity (Tremblay & Small, 2011), whereas in the second study no effect of complexity was observed (McGettigan et al., 2011). Nonetheless, Deschamps and Tremblay (2014) demonstrated that syllabic complexity as defined by the absence or presence of consonant clusters in syllable sequences recruits a broad network of regions within the superior temporal plane.

In the present study, consistent with previous neuroimaging studies (Binder et al., 2000; Demonet et al., 1992; Giraud & Price, 2001; Liebenthal et al., 2005; Okada et al., 2010; Poeppel et al., 2004; Vaden, Muftuler, & Hickok, 2010; Vouloumanos et al., 2001), an effect of phonological complexity as indexed by sonority differences was observed in the left STS. From our whole-brain analysis, we found the locus of activation in the mid portion of the STS. Previous studies have noted clusters of activation within the mid-anterior STS during phonemic/non-phonemic discrimination tasks (Liebenthal et al., 2005), passive listening to speech sounds (Binder et al., 2000) and repetition of words varying in phonemic similarity between items (Vaden et al., 2010). Mid-posterior STS activation has been observed for manipulation of the neighborhood density of words (Okada & Hickok, 2006) or the degree of intelligibility of sentences (Okada et al., 2010). In contrast, reading studies report activation within the posterior STS (Turkeltaub, Eden, Jones, & Zeffiro, 2002). Taken together, these results suggest that subregions within the STS show markedly different patterns of activation depending on the kind of phonological representations being processed (i.e. phonemes, syllables, whole-word) with the
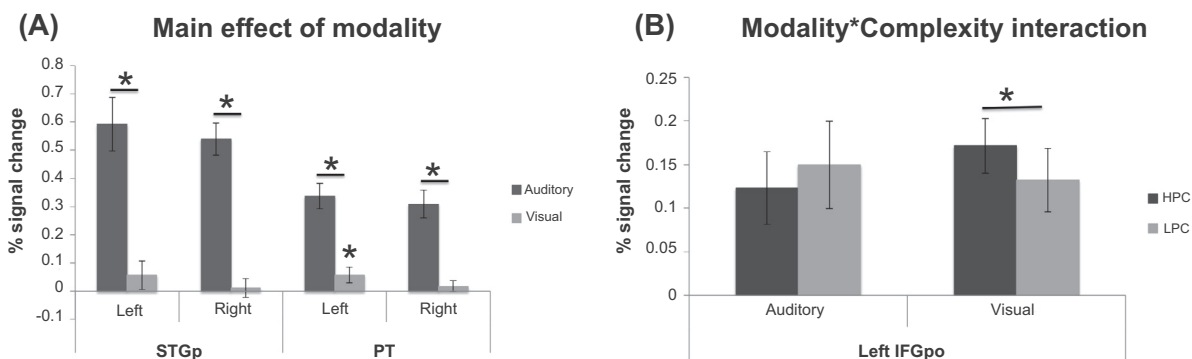

Fig. 2. Brain activity, expressed as a percentage of signal change. Single asterisk indicate a significant difference against zero (one-sample t-test). STGp = posterior superior temporal gyrus; PT = planum temporal; IFGpo = Inferior frontal gyrus pars opercularis.

mid-anterior STS appears sensitive to the processing of phonemic/-syllabic information whereas the mid-posterior STS appears more sensitive to lexical–phonological information. We note that we did not find a complexity or a modality * complexity effect within our STS ROI because the locus of activation observed in the whole-brain analysis is located in the mid STS not the posterior STS. Thus, only the mid STS is sensitive to sonority differences targeting syllabic structure, which is congruent with the hypothesis that different subregions within the STS subserve different phonological processes.

Interestingly, the only region sensitive to phonological complexity in the ROI analysis was the posterior part of the left IFG (IFGpo) and only for print. Despite the absence of an overt response, the orthographic presentation activated different brain areas than passive listening suggesting that the processing of orthographic stimuli results in an obligatory orthographic to phonological transformation or articulatory recoding, a function previously attributed to this region by others (for more details, refer to: Burton, Noll, & Small, 2001; Burton et al., 2000; Burton et al., 2005; Zatorre, Evans, Meyer, & Gjedde, 1992; Zatorre, Meyer, Gjedde, & Evans, 1996). Interestingly, it appears that this transformation/recoding is sensitive to sonority differences at the neural level even during passive reading.

The whole-brain analysis and the ROI analysis also identified regions that were significantly activated for both modalities for pseudowords despite not showing an effect of phonological complexity. These regions are similar to the ones reported in Burton et al. (2005) and Van Atteveldt et al. (2004), suggesting that pseudowords recruit a similar network as words and letters under different task demands. Of particular interest is that the lower bank of the bilateral STG/STS was activated for both auditory and visual pseudowords, an area that has been previously implicated in automatic speech/print convergence processes (i.e. integration) using letters and single speech sounds during passive listening/viewing (Blau, van Atteveldt, Ekkebus, Goebel, & Blomert, 2009; Van Atteveldt et al., 2004). In addition, the ROI analysis revealed that only the left PT was significantly active in both modalities. One potential role of the PT in auditory processing is as a computational hub, disambiguating complex sounds through the isolation of different properties of the acoustic objects (e.g. temporal and spectral information) and matching them to stored phonological spectro-temporal templates (Griffiths & Warren, 2002). However, given that PT was also recruited during the passive reading task, it appears that the PT is involved in accessing cross-modal spectro-temporal profiles. In others words, during passive listening and reading, the PT is involved in accessing stored phonological representations. This is consistent with previous neuroimaging studies that have reported activation within PT under a wide range of experimental paradigms that require access to auditory phonological spectro-temporal templates, such as passive listening to speech sounds (Mazoyer et al., 1993; Petersen, Fox, Posner, Mintun, & Raichle, 1988; Tremblay, Deschamps, & Gracco, 2011; Wise et al., 1991), active listening to speech sounds (Binder, Frost, Hammeke, Rao, & Cox, 1996; Binder et al., 1997; Binder et al., 2000) and reading (Nakada, Fujii, Yoneoka, & Kwee, 2001).

5. Conclusion

In the present study we used sonority difference in pseudowords to evaluate the manner in which differences in phonological properties activate brain regions for spoken and written speech under passive stimulus conditions. The results of the present investigation suggest that while sonority is an important concept in phonological theory, language acquisition and language breakdown, at the neural level, sonority differences alone in pseudowords do not modulate the entire network of regions typically involved in phonological processing. The present results suggest that the phonological properties of speech associated with sonority are insufficient to activate brain areas associated with phonological processing as measured by fMRI. In the visual modality, sonority differences modulated activation within the left IFGpo, suggesting a stronger association between sonority differences and speech processing most likely due to an obligatory decoding of orthographic features into phonological forms. It may be the case that the lack of a sonority effect in the auditory modality might be related to the inherent limitations of fMRI, that is, its poor temporal resolution. Further work using different functional neuroimaging techniques with better temporal resolution, such as electroencephalography (EEG) or magnetoencephalography (MEG) might be of value in capturing sonority effects in speech perception.

Acknowledgements

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bandl.2015.06.008.

References

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., et al. (2007). The English lexicon project. *Behavior Research Methods, 39*(3), 445–459.

Bastiaanse, R., Gilbers, D., & Van Der Linde, K. (1994). Sonority substitutions in Broca's and conduction aphasia. *Journal of Neurolinguistics, 8*(4), 247–255.

Baum, S. R. (2002). Sensitivity to sub-syllabic constituents in brain-damaged patients: Evidence from word games. *Brain and Language, 83*(2), 237–248.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological), 57*(1), 289–300.

Benson, R. R., Richardson, M., Whalen, D. H., & Lai, S. (2006). Phonetic processing areas revealed by sinewave speech and acoustically similar non-speech. *Neuroimage, 31*(1), 342–353.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex, 10*(5), 512–528.

Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience, 17*(1), 353–362.

Binder, J. R., Frost, J. A., Hammeke, T. A., Rao, S. M., & Cox, R. W. (1996). Function of the left planum temporale in auditory and linguistic processing. *Brain, 119*(Pt 4), 1239–1247.

Blau, V., van Atteveldt, N., Ekkebus, M., Goebel, R., & Blomert, L. (2009). Reduced neural integration of letters and speech sounds links phonological and reading deficits in adult Dyslexia. *Current Biology, 19*(6), 503–508.

Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2002). Functional anatomy of intra- and cross-modal lexical tasks. *Neuroimage, 16*(1), 7–22.

Burton, M. W., Locasto, P. C., Krebs-Noble, D., & Gullapalli, R. P. (2005). A systematic investigation of the functional neuroanatomy of auditory and visual phonological processing. *Neuroimage, 26*(3), 647–661.

Burton, M. W., Noll, D. C., & Small, S. L. (2001). The anatomy of auditory word processing: Individual variability. *Brain and Language, 77*(1), 119–131.

Burton, M. W., Small, S. L., & Blumstein, S. E. (2000). The role of segmentation in phonological processing: An fMRI investigation. *Journal of Cognitive Neuroscience, 12*(4), 679–690.

Clements, G. N. (2009). Does sonority have a phonetic basis? Comments on the chapter by Vaux. In R. Eric & C. Charles (Eds.), *Contemporary views on architecture and representations in phonological theory* (pp. 165–175). MIT Press.

Clements, G. N., Kingston, J., & Beckman, M. E. (1990). *The role of the sonority cycle in core syllabification. Papers in Laboratory Phonology*. Cambridge University Press.

Colheart, M. (1981). The MRC psycholinguistics database. *Quaterly Journal of Experimental Psychology, 33A*, 497–505.

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, 29*(3), 162–173.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based Analysis: I. Segmentation and surface reconstruction. *NeuroImage, 9*(2), 179–194.

Demonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., et al. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain, 115*(Pt 6), 1753–1768.

Deschamps, I., & Tremblay, P. (2014). Sequencing at the syllabic and supra-syllabic levels during speech perception: An fMRI study. *Frontiers in Human Neuroscience, 8*, 492.

Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage, 31*(3), 968–980.

Destrieux, C., Fischl, B., Dale, A., & Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage, 53*(1), 1–15.

Dobrovolsky, M., & Katamba, F. (1996). Phonetics: The sounds of a language. In W. D. O'Grady, W. O'Grady, M. Dobrovolsky, & F. Katamba (Eds.), *Contemporary linguistics: An introduction* (pp. 16–58). Longman.

Duvernoy, H. M. (1991). *The human brain: Surface, three-dimensional sectional anatomy with MRI, and blood supply*. New York: Springer-Wien.

Fabre, D., & Bedoin, N. (2003). Sensitivity to sonority for print processing in normal readers and dyslexic children. *Current Psychology Letters. Behaviour, Brain & Cognition, 1*, 10. 2003.

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis: II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage, 9*(2), 195–207.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D. H., et al. (2004). Automatically parcellating the human cerebral cortex. *Cerebral Cortex, 14*(1), 11–22.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage, 15*(4), 870–878.

Gierut, J. A. (2007). Phonological complexity and language learnability. *American Journal of Speech-Language Pathology, 16*(1), 6–17.

Giraud, A. L., & Price, C. J. (2001). The constraints functional neuroimaging places on classical models of auditory word processing. *Journal of Cognitive Neuroscience, 13*(6), 754–765.

Goad, H. (2010). *Phonological processes in child speech: The Oxford handbook of developmental linguistics*. Oxford: Oxford University Press.

Griffiths, T. D., & Warren, J. D. (2002). The planum temporale as a computational hub. *Trends in Neurosciences, 25*(7), 348–353.

Jacquemot, C., Pallier, C., LeBihan, D., Dehaene, S., & Dupoux, E. (2003). Phonological grammar shapes the auditory cortex: A functional magnetic resonance imaging study. *Journal of Neuroscience, 23*(29), 9541–9546.

Kareken, D. A., Lowe, M., Chen, S. H., Lurito, J., & Mathews, V. (2000). Word rhyming as a probe of hemispheric language dominance with functional magnetic resonance imaging. *Neuropsychiatry, Neuropsychology, & Behavioral Neurology, 13*(4), 264–270.

Liebenthal, E., Binder, J. R., Spitzer, S. M., Possing, E. T., & Medler, D. A. (2005). Neural substrates of phonemic perception. *Cerebral Cortex, 15*(10), 1621–1631.

Macchi, L., Schelstraete, M.-A., & Casalis, S. (2014). Word and pseudoword reading in children with specific speech and language impairment. *Research in Developmental Disabilities, 35*(12), 3313–3325.

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., et al. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience, 5*(4), 467–479.

McGettigan, C., Warren, J. E., Eisner, F., Marshall, C. R., Shanmugalingam, P., & Scott, S. K. (2011). Neural correlates of sublexical processing in phonological working memory. *Journal of Cognitive Neuroscience, 23*(4), 961–977.

Mechelli, A., Gorno-Tempini, M. L., & Price, C. J. (2003). Neuroimaging studies of word and pseudoword reading: Consistencies, inconsistencies, and limitations. *Cognitive Neuroscience, Journal of, 15*(2), 260–271.

Miozzo, M., & Buchwald, A. (2013). On the nature of sonority in spoken word production: Evidence from neuropsychology. *Cognition, 128*(3), 287–301.

Morrisette, M. L., Farris, A. W., & Gierut, J. A. (2006). Applications of learnability theory to clinical phonology. *International Journal of Speech-Language Pathology, 8*(3), 207–219.

Nakada, T., Fujii, Y., Yoneoka, Y., & Kwee, I. L. (2001). Planum temporale: Where spoken and written language meet. *European Neurology, 46*(3), 121–125.

Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J.-B. (2005). Valid conjunction inference with the minimum statistic. *NeuroImage, 25*(3), 653–660.

Obleser, J., Zimmermann, J., Van Meter, J., & Rauschecker, J. P. (2007). Multiple stages of auditory speech perception reflected in event-related FMRI. *Cerebral Cortex, 17*(10), 2251–2257.

Okada, K., & Hickok, G. (2006). Identification of lexical–phonological networks in the superior temporal sulcus using functional magnetic resonance imaging. *NeuroReport, 17*(12), 1293–1296.

Okada, K., Rong, F., Venezia, J., Matchin, W., Hsieh, I. H., Saberi, K., et al. (2010). Hierarchical organization of human auditory cortex: Evidence from acoustic invariance in the response to intelligible speech. *Cerebral Cortex, 20*(10), 2486–2495.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature, 331*(6157), 585–589.

Poeppel, D., Guillemin, A., Thompson, J., Fritz, J., Bavelier, D., & Braun, A. R. (2004). Auditory lexical decision, categorical perception, and FM direction discrimination differentially engage left and right auditory cortex. *Neuropsychologia, 42*(2), 183–200.

Poldrack, R., Temple, E., Protopapas, A., Nagarajan, S., Tallal, P., Merzenich, M., et al. (2001). Relations between the Neural Bases of Dynamic Auditory Processing and Phonological Processing: Evidence from fMRI. *Journal of Cognitive Neuroscience, 13*(5), 687–697.

Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage, 62*(2), 816–847.

Rimol, L. M., Specht, K., & Hugdahl, K. (2006). Controlling for individual differences in fMRI brain activation to tones, syllables, and words. *Neuroimage, 30*(2), 554–562.

Romani, C., & Calabrese, A. (1998). Syllabic constraints in the phonological errors of an aphasic patient. *Brain and Language, 64*(1), 83–121.

Romani, C., & Galluzzi, C. (2005). Effects of syllabic complexity in predicting accuracy of repetition and direction of errors in patients with articulatory and phonological difficulties. *Cognitive Neuropsychology, 22*(7), 817–850.

Saad, Z. S., Reynolds, R. C., Argall, B. D., Japee, S., & Cox, R. W. (2004). SUMA: An interface for surface-based intra- and inter-subject analysis with AFNI. In *Paper presented at the proceedings of the IEEE international symposium on biomedical imaging*.

Saad, Z. S., Glen, D. R., Chen, G., Beauchamp, M. S., Desai, R., & Cox, R. W. (2009). A new method for improving functional-to-structural MRI alignment using local Pearson correlation. *Neuroimage, 44*(3), 839–848.

Shaul, S., Arzouan, Y., & Goldstein, A. (2012). Brain activity while reading words and pseudo-words: A comparison between dyslexic and fluent readers. *International Journal of Psychophysiology, 84*(3), 270–276.

Simos, P. G., Breier, J. I., Fletcher, J. M., Foorman, B. R., Castillo, E. M., & Papanicolaou, A. C. (2002). Brain mechanisms for reading words and pseudowords: An integrated approach. *Cerebral Cortex, 12*(3), 297–305.

Solso, R. L., & Juel, C. L. (1980). Positional frequency and versatility of bigrams for two-through nine-letter English words. *Behavior Research Methods and Instrumentation, 12*(3), 297–343.

Sperbeck, M., & Strange, W. (2010). The perception of complex onsets in English: Universal markedness? *University of Pennsylvania Working Papers in Linguistics, 16*(11). Article 22.

Steriade, D. (1990). Gestures and autosegments. In J. Kingston & M. Beckman (Eds.), *Papers in laboratory phonology I* (pp. 382–397). Cambridge: Cambridge University Press.

Tremblay, P., Baroni, M., & Hasson, U. (2012). Processing of speech and non-speech sounds in the supratemporal plane: Auditory input preference does not predict sensitivity to statistical structure. *Neuroimage, 66C*, 318–332.

Tremblay, P., Deschamps, I., & Gracco, V. L. (2011). Regional heterogeneity in the processing and the production of speech in the human planum temporale. *Cortex*.

Tremblay, P., & Small, S. L. (2011). On the context-dependent nature of the contribution of the ventral premotor cortex to speech perception. *Neuroimage, 57*(4), 1561–1571.

Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. *Neuroimage, 16*(3 Pt 1), 765–780.

Vaden, K. I., Jr., Muftuler, L. T., & Hickok, G. (2010). Phonological repetition-suppression in bilateral superior temporal sulci. *Neuroimage, 49*(1), 1018–1023.

Van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron, 43*(2), 271–282.

Vouloumanos, A., Kiehl, K. A., Werker, J. F., & Liddle, P. F. (2001). Detection of sounds in the auditory stream: Event-related fMRI evidence for differential activation to speech and nonspeech. *Journal of Cognitive Neuroscience, 13*(7), 994–1005.

Wise, R., Chollet, F., Hadar, U., Friston, K., Hoffner, E., & Frackowiak, R. (1991). Distribution of cortical neural networks involved in word comprehension and word retrieval. *Brain, 114*(Pt 4), 1803–1817.

Xiao, Z., Zhang, J. X., Wang, X., Wu, R., Hu, X., Weng, X., et al. (2005). Differential activity in left inferior frontal gyrus for pseudowords and real words: An event-related fMRI study on auditory lexical decision. *Human Brain Mapping, 25*(2), 212–221.

Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science, 256*(5058), 846–849.

Zatorre, R. J., Meyer, E., Gjedde, A., & Evans, A. C. (1996). PET studies of phonetic processing of speech: Review, replication, and reanalysis. *Cerebral Cortex, 6*(1), 21–30.

Zec, D. (1995). Sonority constraints on syllable structure. *Phonology, 12*(01), 85–129.